Navy Personnel Research and Development Center

San Diego, CA 92152-6800 TN 89-6 October 1988

Brain Activity During Tactical Decision-making: IV. Event-related Potentials as Indices of Selective Attention and Cognitive Workload

NPRDC TN 89-6 October 1988

BRAIN ACTIVITY DURING TACTICAL DECISION-MAKING: IV. EVENT-RELATED POTENTIALS AS INDICES OF SELECTIVE ATTENTION AND COGNITIVE WORKLOAD

Mark H. Blankenship

Leonard J. Trejo

Gregory W. Lewis

Approved and released by
J.C. McLachlan
Director, Training Systems

Navy Personnel Research and Development Center
San Diego, California 92152-6800

FOREWORD

This report is the fourth in a series of reports examining the feasibility of using neuroelectric signals to predict decision-making of combat system operators under varying workloads. The first report (HFOSL Tech. Note 71-86-6) identified assumptions underlying this approach to the study of decision-making. The second report (NPRDC TN 88-12) provided detailed analyses of the physiological changes in brain activity that occur in response to an irrelevant visual probe as cognitive workload increases in a combat system simulation.

The third report (in press) describes relationships between physiological brain activity and a combat system simulation and on-job performance variates. This report provides analyses of the physiological changes in brain activity that occur in a dual-task paradigm as cognitive workload increases in a combat systems simulation.

Research described in this report is being performed under the work units 521-804-042.03.2 (Future Technologies-Biopsychometrics) sponsored by the Office of Naval Technology and 44-521-080-203 (Biopsychometric Assessment) sponsored by Headquarters, Marine Corps (MA).

J.C. McLACHLAN
Director, Training Systems

SUMMARY

Problem

The demands of modern combat systems have the potential for exceeding the capacity of the human to accurately process information, especially during times of great stress. The capacity of the human to perceive, integrate, remember, and use information may be challenged when the individual is flying aircraft, monitoring radar and sonar displays, or operating electronic warfare systems. Exceeding the capacity of the human operator in such situations may impair decision-making and could result in costly tactical errors.

Although much is being done to improve the reliability of combat systems, not enough is being done to improve the system operators. For these reasons, the most unpredictable element in combat systems is the human operator. Years of personnel testing have not eliminated this unpredictability. In part, this is because traditional testing methods tend to measure what a person *knows rather than how a person thinks and processes information*.

This research is driven by the Navy's need for better methods of assessing combat system operators, particularly for predicting the ability of operators to continue to make accurate decisions under heavy workloads.

Objective

This report, the fourth in a series of reports concerned with the use of neuroelectric signals to predict the decision-making performance of combat system operators, provides detailed analyses of the neuroelectric changes that occur as workload increases in a combat system simulation.

Approach

We presented visual and auditory oddball stimuli to 65 male U. S. Marine Corps volunteers while they were engaged in an anti-air combat simulation (AIRDEF). Each subject performed the simulation at three levels of workload, which were defined in terms of the number of targets that appeared on the radar display. The visual oddball stimuli were diffuse flashes of light with a duration of 16 milliseconds, presented at irregular intervals. The auditory oddball stimuli were tone bursts with a duration of 10 milliseconds, presented at irregular intervals.

Under each condition, event-related potentials (ERPs), which are physiological measures of brain function resulting from the oddball stimuli, were recorded from eight electrodes covering the frontal, temporal, parietal, and occipital areas of the scalp. One electrode was placed above the eye to monitor eye movements. A nose electrode was the reference for all recordings. Each single ERP was first analog-filtered (3 dB bandwidth 0.1-100 Hz), then sampled at 128 Hz, digitized, and stored by a computer. Signal average waveforms were computed from 10 artifact-free ERPs for each condition. (Each point in the signal average waveform is the time-indexed average of the 10 single ERPs.) The waveforms were digitally filtered (0.5-25 Hz) and divided into 10 adjacent, non-overlapping time windows approximately 50 ms wide that spanned the range between 42 and 511 ms after stimulus onset. The root-mean-square (RMS) value of the waveform was computed in each window. These RMS values were used as dependent variables in a repeated measures analyses of variance (within-subjects factors were *workload*, *electrode position*, and *window latency*) and a repeated measures analysis of covariance (within-subjects factors were the same as above and the covariate was RMS amplitude of the N1 component).

Results

Results demonstrated that a specific ERP RMS amplitude measure peaking between 300-450 ms post-stimulus at two electrode sites (P3 and F4) was sensitive to workload changes in the air defense simulation. This finding agrees with studies in the literature that have used a similar time window in the ERP to measure workload. The results were enhanced by taking into consideration an early ERP amplitude measure, peaking between 90-140 ms post-stimulus, which corresponds to an ERP component that has been related to selective attentional factors.

Implications

This study indicates that ERPs can be used to monitor selective attention and other cognitive processes associated with performance on a simulated air defense task. This evidence suggests that ERPs may also be used in the on-line monitoring of cognitive resources. This monitoring could aid in tactical decision-making during critical task operations.

CONTENTS

LIST OF TABLES

INTRODUCTION

The Navy depends heavily on cognitive tests, such as the Armed Services Vocational Aptitude Battery, to evaluate personnel. The results of such testing can predict academic performance reasonably well (Fishman, 1958), but are less effective in predicting on-job performance (Ghiselli, 1966, 1973). There is a need for new kinds of assessment procedures that will supplement the information derived from written tests and provide an improved understanding and a more complete assessment of the unique capability of each individual (Lewis & Sorenson, 1987). In an attempt to better understand the human as an integrator and decision maker in operating systems, recordings of event-related brain potentials (ERPs) have been employed to assess individual brain processes and their relationship to differences in on-job performance and decision making (Lewis, 1979; Lewis, 1983b; Trejo, Lewis, & Blankenship, 1987).

The emphasis of the present research is on using brain activity to assess and ultimately to predict cognitive resources and decision-making performance of combat system operators. A previous report (Trejo, 1986) discussed the background for this research, including assumptions and hypotheses. An air defense radar simulation (AIRDEF) (Kelly, Greitzer, & Hershman, 1981) was adapted and used as a tool for assessing brain activity and behavior during decision making under various dual-tasking situations.

Research has shown that features of the ERP are sensitive to variations in task demands for cognitive resources (for reviews, see Donchin, Kramer, & Wickens, 1982; Gopher & Donchin, 1986). Relationships between ERPs and resource demands are easily observed in experiments structured around a dual-task situation. This *dual-task paradigm* requires a subject to perform two tasks concurrently (Brown, 1978; Rolfe, 1971). One of the tasks is designated as the primary task and the other the secondary task. The technique is useful "for the study of individual differences in processing resources when the additional [secondary] task presents discrete stimuli which impose [a] constant load, is carried out at a forced pace and competes with the primary task for processing resources" (Brown, 1978, p. 221). The major findings reviewed suggest that the P300 component[1] of the ERP elicited by secondary task stimuli is affected by the difficulty of the primary task. Namely, secondary-task P300 amplitude decreases and latency increases as a function of primary task difficulty.

In a recent study (Sirevaag, Kramer, Coles, & Donchin, in press), ERPs were recorded to *both* the primary and secondary tasks. Results for secondary task ERPs agreed with earlier findings; however, the amplitude of P300 elicited by primary task stimuli *increased* with primary task difficulty. These results support a model of the human as a resource-limited processor (Norman & Bobrow, 1975) in the following way: The *resource allocation* model predicts that if a task demands more cognitive resources than an individual has available, a decrement in the performance of that task will occur in relation to the ratio of required resources to total resources available. Likewise, if an individual is performing two concurrent tasks that require more resources than the individual has available, a performance decrement will occur in one or both of the tasks depending on task priorities. Other psychophysiological studies have used irrelevant probe stimuli in attempts to study cognitive processes (Papanicolaou & Johnstone, 1984; Trejo et al., 1987).

Several ERP studies have been conducted using the dual-task paradigm to evaluate human cognitive capacity. (The term capacity refers here to resources which subserve processes such as attention, short-term memory, and decision making.) These studies have provided information about the physiological components that are related to quality of decision making, stimulus evaluation, and information processing in the brain. In many of these studies the primary task included complex behavioral performance, while the secondary task consisted of an *oddball* paradigm[2] (e.g., Defayolle, Dinard, & Gentil, 1971; Isreal, Chesney, Wickens, & Donchin, 1980; Karis, Coles, & Donchin, 1984; Kramer, Wickens, & Donchin, 1986; Lindholm, Cheatham, & Koriath, 1984). In some of the studies the oddball paradigm was combined with a compensatory tracking task in an attempt to index cognitive workload.

---

[1]Sutton, Braren, Zubin, & John (1965) were the first to report that the probability of a stimulus occurring had systematic effects on a large positive-going component in the ERP about 300 ms after the presentation of the stimulus. This component came to be known as P300 because of its polarity and latency. The Sutton et al. experiment indicated that there was an inverse relationship between stimulus probability and the P300 component. Much research has been done on the P300 component since the original Sutton et al. study. There are still many controversies as to the exact nature of P300 and how it is defined and measured (Duncan-Johnson & Donchin, 1979; Johnson, 1978, 1986; Johnson & Donchin, 1985; Pritchard, 1981).

[2]The *oddball* paradigm (Gopher & Donchin, 1986) consists of presenting two stimuli with different probabilities of occurrence that differ with respect to some physical property (i.e., intensity). The stimuli are usually presented on a .75/.25 basis with the stimuli occurring most often being referred to as the *frequent* stimuli and the stimuli occurring least often as the *rare* stimuli.

In the first study (Wickens, Isreal, & Donchin, 1977), subjects performed a tracking task in which difficulty was varied by increasing the number of tracking dimensions from one to two (horizontal and vertical). An auditory oddball task served as the secondary task. The results showed that the amplitude of P300 elicited by the rare auditory stimuli decreased dramatically with the introduction of the tracking task, but failed to change when tracking difficulty was increased from one to two dimensions. In the second study (Isreal, Chesney, Wickens, & Donchin, 1980), the difficulty of the tracking (primary) task was varied by changing the limits of variation in the velocity of the cursor. Again, amplitude of P300 elicited by rare auditory stimuli decreased when the tracking task was introduced, but did not change significantly with changes in the difficulty of the primary task. These results showed that P300 amplitude may have a limited dynamic range when it serves as an index of mental workload in the dual-task paradigm. Later work suggests that this limit in dynamic range results not from a connection between P300 and motor demands of a task, but from a connection between P300 and perceptual processing demands of the task, which did not change significantly with task difficulty (Donchin, Kramer, & Wickens, 1986; Gopher & Donchin, 1986; Isreal, Wickens, Chesney, & Donchin, 1980; Wickens, 1980). A multiple resource explanation was given to explain the experimental data.

To test the connection between P300 and perceptual processing demands, Isreal, Wickens, Chesney, and Donchin (1980) manipulated the perceptual processing demands of an air-traffic controller's display (modeled after Heffley, Wickens, & Donchin, 1978). The task consisted of a subject monitoring squares (targets) and triangles (noise) traversing an oscilloscope. Level of difficulty was manipulated by varying the number of objects to be monitored. In each condition, subjects were required to monitor for course changes in the squares or count visual intensity changes in the squares. When subjects monitored for course changes, they were required to perform an auditory oddball task concurrently. Results indicated that P300 amplitude decreased monotonically to the auditory stimuli when subjects monitored for course changes. When subjects monitored visual intensifications, P300 decreased with the introduction of the task, but failed to decrease with increases in the number of targets. It appears that monitoring the number of targets alone was not enough. The change in P300 amplitude had to interact with either the task (course change or intensity change) and/or the modality of the secondary task (auditory vs. visual).

A study conducted by Kramer, Wickens, Vanasse, Heffley, and Donchin (1981) using a pursuit step tracking task--a task where subjects must keep a cursor superimposed on a moving target--confirmed the results of the Isreal, Wickens, Chesney, and Donchin (1980) study. A major note of difference, however, was the inclusion of a condition in which subjects engaged the task at varying levels of difficulty, but were not required to attend to a visual probe. Results of the P300 amplitude to this "irrelevant" visual probe were the same as the P300 amplitude to the conditions when the flash was attended to (Kramer et al., 1981). These data indicated that processing resources and resource allocation could be measured in the absence of a secondary task altogether (Kramer et al., 1981).

Based on the experimental evidence suggesting that task irrelevant stimuli could be used in the assessment of mental workload and resource allocation (Kramer et al., 1981; Papanicolaou & Johnstone, 1984), Trejo et al. (1987) conducted a study using an anti-air warfare simulation as a primary task, while collecting ERPs to an irrelevant visual probe. The simulation's processing requirements were similar to those of the air traffic controller simulation used by Isreal, Wickens, Chesney, and Donchin (1980). Subjects monitored a radar screen for incoming enemy missiles and attempted to defend their ship in the center of the display by launching missiles at them. Level of difficulty was manipulated by varying the number of incoming missiles. The simulation required more decision-making processes by the subject than those required by the air-traffic controller task.

Results of the study indicated that ERP amplitudes for specific recording sites and time (latency) windows were more sensitive to task difficulty than others. As with previous studies, there were significant amplitude changes in these sites (frontal, parietal, and occipital) and windows (48-ms windows centered at 127 ms, 229 ms, and 330 ms) with the introduction of the task, but increases in task difficulty failed to further attenuate the ERP component amplitudes.

P300 has not been the only ERP component used in the study of workload and attentional capacities. Several investigators have conducted experiments with oddball and other paradigms in selective attention research (see Näätänen & Picton, 1987 for a comprehensive review). The primary component of interest in these studies has been labeled the N1 component.[3] Under single-tasking conditions, the N1 component is enhanced by relevant stimulus processing in much the same way that P300 is enhanced.

[3]This component is labeled N1 because of its polarity (negative-going wave) and peak latency (100 ms after stimulus presentation). As with P300, there is no agreed upon meaning of the actual processes involved in the generation of N1 (Näätänen & Picton, 1987).

Several investigators have reported that an early component of the ERP is related to early selective attention processes or "stimulus set" responses (Näätänen & Picton, 1987). Depending on the modality of the stimulus (visual vs. auditory), this component is negative and has a peak latency between 100 and 140 ms. This is the N1 component. In selective attention studies where subjects have been required to pay attention to a particular stimulus and ignore other stimuli being presented simultaneously, the amplitude of the N1 component has been enhanced, but not so when the stimulus was ignored (see Näätänen & Picton, 1987, for a comprehensive review). The N1 component has not been found to be related to processing resources or capacity, but rather is enhanced simply by attention to a stimulus. Its resource pool may be different from the resource pool that P300 reflects.

If P300 is, in fact, invoked by task-relevant stimuli that require the subject's processing, and N1 is enhanced by attending to a stimulus set, then, intuitively, attentional factors and processing factors should be related. If we can further assume that P300 amplitude is an index of processing resources while N1 amplitude is insensitive to those resources per se, but indexes attentional gating instead, then the physiological data could be used to assess whether or not a subject has, in fact, attended to the stimulus. This additional information should increase prediction and assessment of workload-related processing if later components, namely P300, assess processing resources.

In addition to the dual-task paradigm, other methods of neuroelectric (and neuromagnetic) recording have also been used to evaluate human cognitive capacity. Several studies conducted at the Navy Personnel Research and Development Center in San Diego have shown that features of ERPs elicited by simple sensory stimuli in the absense of response requirements are related to global features of human performance (Lewis, 1979, 1983a, 1983b, 1984; Lewis & Froning, 1981; Lewis & Rimland, 1979; Lewis & Rimland, 1980; Lewis, Rimland, & Callaway, 1977; Lewis, Trejo, Blackburn, & Blankenship, 1986). These studies, which have emphasized ERP energy measures rather than component amplitudes or latencies, indicate a potential for the use of psychophysiological methods in the assessment of personnel (Lewis & Sorenson, 1987).

The above studies have shown that ERP components such as N1 and P300 may aid in the assessment of individual differences in cognitive resources. Researchers have tended to pick specific components in the ERP such as N1 or P300 and limit analyses to their relationships to psychological processes. This report deals with the potential relationship between N1 and P300 and their interaction with cognitive workload.

We now report an analysis of workload-related changes in the neuroelectric correlates of visual and auditory oddball stimuli recorded during the performance of AIRDEF by a group of U. S. Marines.

METHODS

Subjects

Sixty-five male volunteers from the U. S. Marine Corps attending the non-commissioned officers (NCO) Leadership School at Camp Pendleton participated in the study. Subjects were randomly assigned to one of six groups in order to counterbalance the order of task difficulty. All subjects who received the same order of difficulty presentations in the experimental task constituted a "group." The size of each group follows: Group 1, 10; Group 2, 11; Group 3, 12; Group 4, 11; Group 5, 11; and Group 6, 10.

All subjects signed a consent form that assured the confidentiality of their data in accordance with Department of Defense regulations. Subjects were given a copy of the Privacy Act of 1974 (Public Law 93-579), which implements 5U.S.C552A (Privacy Act of 1974) and DOD Directive 5400.11 (Personal Privacy and Rights of Individuals). Subjects were also given a copy of the consent form. After a subject read the two forms, the experimenter answered any questions the subject had concerning the experiment. Each subject then signed the consent form and filled out the biographical questionnaire. The questionnaire results (see Table 1) indicated that the subjects were mostly right-handed Caucasians in their early twenties. Most subjects reported that they were not on medication, were not tired or drowsy, and had little sight or hearing difficulty. About half of the subjects smoked or chewed tobacco. Subjects had the right to withdraw at any time during the experiment if they so desired.

Air Defense Simulation (AIRDEF)

AIRDEF simulates anti-air warfare as fought from a ship. A schematic diagram of the simulation in progress is shown in Figure 1 (after Kelly, Greitzer, & Hershman, 1981; Trejo, 1986; Trejo et al., 1987). The major

Table 1

Summary of Biographical Data

Age (years)		Height (inches)		Weight (pounds)	
Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
23.3	3.0	70.53	2.73	172.41	20.48
Race				**Handedness**	
Caucasian	Black	Hispanic	Other	Left	Right
77%	14%	4.5%	4.5%	8%	92%
Eye Color			**Hair Color**		
Brown	Blue	Green	Brown	Black	Blond
52%	32%	16%	59%	19%	22%
Arousal				**Languages Spoken**	
Tired	Drowsy	Awake	Alert	English Only	English and Other
9%	24%	49%	18%	78%	22%
Tobacco Use		**Wear Glasses/Contact Lenses**		**Difficulty Hearing**	
Use	No Use	No	Yes	No	Yes
52%	48%	66%	34%	95%	5%

components of the display are highlighted. A blue radar sweep and blue range indicators detected incoming hostile missiles and tracked their position with orange blips. Missiles appeared unpredictably from 360 degrees with at least 4 degrees between them. Each incoming missile appearing on the screen was assigned a random two-digit tracking number. Incoming missiles traveled toward the subject's ship which was at the center of the display, and was represented by a "+". The missiles traveled at three different speeds--slow, medium, and fast--with each missile speed constituting one-third of the missile total. The speed of any given missile was denoted by the spacing between its tracking blips; the faster the missile, the farther the blips were spaced apart.

The outer most circle represented the range of the ship's radar. The inner circle represented the maximum range of the ship's weapons. The subject's task was to avoid getting hit by the incoming missiles. A subject launched his missiles by moving a cursor over the desired tracking number with a mouse (an electronic hand-held cursor control device), and then pressing a button.

The ship's missiles were shown as green blips which traveled straight toward the incoming missiles at the speed of the fastest incoming missiles. If a subject launched too early on an incoming missile, his missile reached its maximum range before the incoming missile arrived and it "splashed." There was no penalty for a splash. On the other hand, if a subject failed to launch on an incoming missile, the incoming missile hit the ship. This was denoted by marking the track with an orange line. A "hit" resulted in a 12-point penalty. Only one missile could be launched on an incoming target at a time. If a subject attempted to launch on a target that already had a missile "in flight," a 2-point penalty resulted. A subject gained points based on a calculation of the average kill range for the incoming missiles. When a subject killed an incoming missile, its track was marked with a green line.

The level of difficulty was controlled by varying the frequency of incoming missiles. For the three difficulty levels employed in the present experiment (low, medium, and high), missile frequencies were approximately 0, 3, and 9 missiles per minute. The speeds of the individual missiles (fast, medium, and slow) did not vary with task difficulty.

Stimulus Presentation

Sixty visual and 60 auditory oddball stimuli were presented during each level of AIRDEF difficulty. Within each modality, the probabilities of the frequent and rare stimuli were .75 and .25, respectively. Thus, 45 frequent and 15 rare stimuli occurred for each modality. Only one stimulus occurred at a time. The mean inter-stimulus interval (ISI) was 2 seconds, with a range of 1.5 to 2.5 seconds.

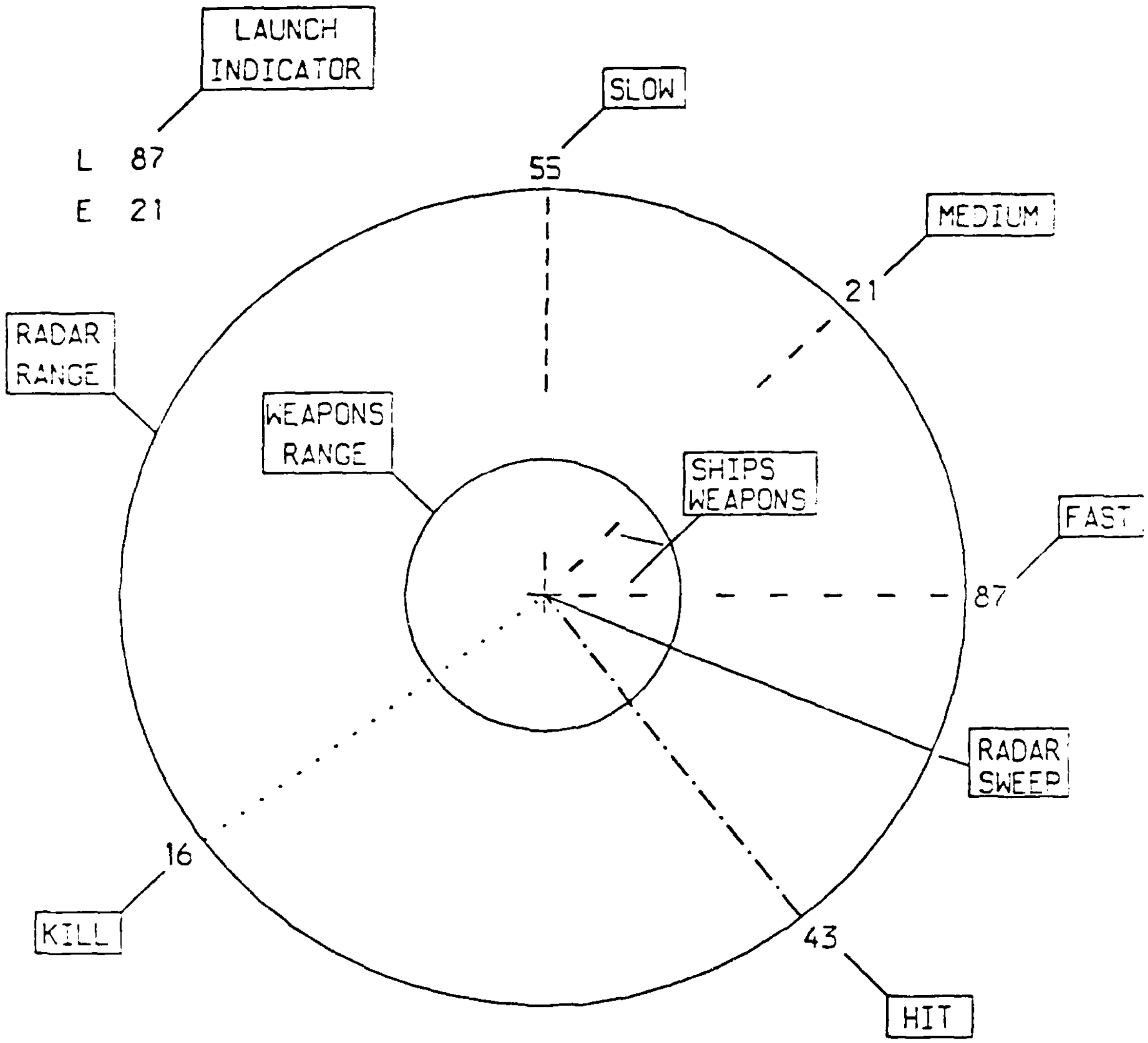

Figure 1. Schematic diagram of the air defense radar simulation. The cross in the center shows the location of the subject's ship and the inner circle marks the maximum weapons range. The outer circle marks the maximum detection range for the radar (these components are color-coded in blue). Track numbers along the outer circle identify incoming targets. Blips represent the status of incoming targets and outgoing weapons (these are color-coded orange and green, respectively). The letters in the upper left corner of the display provide feedback to the subject after each attempted weapon firing. *L* indicates a valid launch on the track number displayed to its right and *E* indicates an error launch, which occurs if the requested track (in this instance, track "21") already has a weapon in flight. When the subject "kills" an incoming target, the blips for that target are marked with a solid green line (shown here as a dotted line). When an incoming target "hits" the subject's ship, that target's blips are marked with a solid orange line (shown here as dashed-dot line). For more details see Trejo et al. (1987).

Auditory Oddball Stimuli

The auditory oddball stimuli (45 frequent and 15 rare for each difficulty level) were synthesized pure tone bursts presented binaurally through a loudspeaker placed one meter directly in front of the subject, just above the computer monitor. Both the frequent and rare stimuli had a duration of 10 ms and an intensity of 75 dB(A) sound pressure level (SPL) (Brüel & Kjaer Impulse Sound Level Meter, Model 2209[4]). The frequent stimulus frequency was 1,500 Hz and the rare stimulus frequency was 750 Hz.

Visual Oddball Stimuli

The visual oddball stimuli (45 frequent and 15 rare for each difficulty level) were spatially uniform flashes presented on a rectangular area of the computer monitor used for the AIRDEF display. These stimuli were background flashes which did not interfere with the presentation of the task data on the screen. Both the frequent and rare stimuli had a duration of 16 ms. This time was limited by the computer graphics processor (screen refresh rate). When viewed from the subject's distance of 45 cm, the stimuli subtended 37° (visual angle) vertically and 49° horizontally. The frequent stimulus luminance was 21.6 nits and the rare stimulus luminance was 5.1 nits (Tektronix Model J16 Digital Photometer).

Data Acquisition

The recording equipment consisted of an array of tin electrodes embedded in a nylon cap (Electro-Cap International), a set of amplifiers (Grass Model 12A5, Neurodata Acquisition System), and a computer (Masscomp Model MC-5500 field-portable) programmed to digitize and record neuroelectric signals. The ERPs were bandpass filtered (analog, 3 dB corner frequencies at 0.1 and 100 Hz), amplified (20,000 gain), sampled at 128 Hz, and digitized. The digitizer range was -250 to +250 µV, with a digitization step size of 0.122 µV. All ERP data were acquired for a period of 125 ms pre-stimulus and 875 ms post-stimulus for a total of one second, and stored as single epochs. Recordings were made from eight sites: F3 and F4 (frontals), T3 and T4 (temporals), P3 and P4 (parietals), and O1 and O2 (occipitals), referred to nose (International 10-20 System; Jasper, 1958). Recordings were also made from site Fp2 (right frontal pole) to monitor eye blinks and large eye movements. Subject ground was at Fz. All the data were collected with a constant background auditory white noise of 60 dB(A) SPL and an ambient illumination of 5.38 lux.

On-line Artifact Rejection

All recorded sites were monitored on line with an eight-channel oscilloscope (Tektronix, Model 5103N). Two two-channel storage oscilloscopes were used for on-line rejection of bad epochs (Tektronix, Model 336). The criterion for epoch rejection was based on the ERPs recorded at site Fp2. ERPs that showed a transient signal with a baseline-to-peak amplitude greater than 50 µV during 500 ms of the post-stimulus period were rejected. If an ERP at Fp2 was rejected, then the corresponding ERPs at every recording site were rejected. Rejected ERPs were repeated to ensure the collection of the desired number of epochs for each stimulus type.

Procedure

After filling out forms and the questionnaire, subjects were fitted with an electrode helmet. Electrode impedances were usually within 1-2 Kohms and did not exceed 5 Kohms. The subject was then seated approximately 45 cm away from the computer monitor in a non-reclining armchair. At this point the subject was given instructions for AIRDEF by the experimenter. The experimenter also demonstrated the use of the mouse to the subject. Subjects then received practice in using the mouse. The operation of the mouse for AIRDEF performance is a relatively simple psychomotor task and competence was attained within a minute of practice. Subjects then engaged in a practice session of AIRDEF that included the three difficulty levels used in the experiment; low load (baseline), medium load, and high load. Each practice level lasted one minute.

The low load condition consisted of a subject moving a cursor over pseudo-randomly appearing blocks where tracking numbers usually appeared. Subjects simulated launching on these blocks, but no missiles were incoming or outgoing. This condition was used as a baseline, and controlled for psychomotor effects on the physiological data.

---

[4]Identification of the equipment is for documentation only and does not imply endorsement

This condition lasted 6 minutes. The medium and high load conditions also lasted 6 minutes and involved the incoming and outgoing missiles described earlier.

During the practice session both visual and auditory oddball stimuli were presented. The experimenter told the subjects that these stimuli would occur while they were performing the simulation and that they would be instructed by the computer on which stimulus they should attend to, either rare visual or rare auditory. Subjects were told to keep a mental count of the attended rare stimulus and that they would be called upon to report this number at the end of each engagement. Subjects were instructed to perform both the AIRDEF simulation (primary) and the counting (secondary) task to the best of their ability. After the experimenter had given the subject proper instruction, he answered any questions that the subject had about operating the mouse or performing the AIRDEF simulation. The electrode impedances were rechecked before the simulation began. Subjects were then instructed to relax their jaw and face muscles as much as possible to minimize muscle artifacts while recording; the experimenter also asked the subjects to minimize eyeblinks.

At the beginning of each trial, the subject was instructed by a visual display on the computer monitor about which rare stimulus to attend to. There was an approximate one-minute break between each trial of the AIRDEF task (see Experimental Design subsection for more detail).

Digital Filtering

After the data had been collected and stored, they were detrended (mean and linear slope removed) and digitally filtered (windowed finite impulse response filter, 127 coefficients, corner frequencies at 0.5 and 25 Hz, Hamming window) in the same manner as the Trejo et al. (1987) study.

Off-line Artifact Rejection

In order to further remove eye blink and/or eye movement artifacts that may have failed to be rejected by the experimenter on line, the on-line rejection criterion was re-applied to the detrended and digitally filtered ERPs; however, the time range for rejection was restricted to the post-stimulus time period of 43 to 512 ms, because this was the window to be analyzed.

Selection of ERPs for Analysis

As a result of the off-line artifact rejection procedure, the number of artifact-free epochs varied between subjects. Three criteria were then applied to select artifact-free epochs for analysis. First, the number of epochs for each stimulus type (visual and auditory rare) had to be equal across the three difficulty levels of AIRDEF (low load, medium load, and high load). Second, the number of epochs had to be equal across all subjects. Third, the number of epochs had to be one that would retain a majority of the subject sample for analysis. Based on those criteria, all 65 subjects had at least 10 artifact-free ERPs for each stimulus type in the three difficulty levels of AIRDEF.

Because workload (target frequency) in AIRDEF rises at the beginning of the engagement and falls near the end, ERPs recorded near the beginning or the end may not show workload-related effects. ERPs near the beginning and near the end of an AIRDEF trial were excluded from analysis whenever more than 10 artifact-free ERPs were available. This procedure resulted in retaining 10 ERPs for each stimulus type that were generally from the middle of an AIRDEF engagement where workload is nearly constant. This follows the procedure used by Trejo et al. (1987) in dealing with unequal numbers of artifact-free ERPs.

Signal Processing

From the 10 ERPs selected for each subject for both visual and auditory rare stimuli, an average waveform was computed, as described in Trejo et al. (1987).

The average wave forms for each subject (representing the 43- to 512-ms post-stimulus window) were then used to compute root mean square (RMS) amplitudes for non-overlapping windows in the post-stimulus time period. The rationale for using RMS amplitudes stemmed from an inspection of the average wave forms and previous experimental findings (Trejo et al., 1987). Basically, the morphology of the average ERP wave forms varied between individuals. This made component analyses difficult and could have resulted in a significant loss of data. By using RMS amplitudes from a standard set of time windows, an unambiguous dependent amplitude measure could be computed for each individual for each time window.

For each time window, the RMS amplitude was computed in µV units. These RMS amplitude values were used as the dependent measure in statistical analyses. The RMS amplitude measure has been used and discussed elsewhere (Callaway, 1975; Callaway, Halliday, & Herning, 1983; Lewis, 1983b; Lewis & Sorenson, 1987; Trejo et al., 1987).

Experimental Design

A counterbalanced repeated measures factorial design was used. Subjects were randomly assigned to one of the six subject groups. They then engaged in the AIRDEF simulation at three levels of difficulty (low, medium, and high). The design was completely counterbalanced with regards to the order of difficulty presentation. Each level of difficulty was presented twice. Within each set of two presentations for each level of difficulty, both visual and auditory oddball stimuli occurred. However, subjects were instructed to pay attention to only one type of stimulus for any one engagement. There was a .5 probability that the subject would be instructed to attend to either the visual or the auditory stimulus in the first trial of f difficulty. On the second trial of the set for a particular difficulty level, the subject was instructed to attend to the other stimulus modality.

RESULTS

AIRDEF Behavioral Performance

Behavioral performance on AIRDEF was first described by Kelly et al. (1981) as being a linear composite of AIRDEF behavioral performance variables that, taken together, comprised an overall skill rating (SR);

$$SR = \left[(5 \cdot \textit{average kill range}) - (12 \cdot \textit{total hits}) - (2 \cdot \textit{number of inflights}) \right] \qquad (1)$$

This composite was later normalized to the number of incoming targets (after Trejo et al., 1987) and used as the normalized skill rating (NSR);

$$NSR = \left[(5 \cdot \textit{average kill range}) - (12 \cdot \frac{\textit{total hits}}{\textit{number of targets}}) - (2 \cdot \frac{\textit{total inflights}}{\textit{number of targets}}) \right] \qquad (2)$$

The Trejo et al. (1987) normalization procedure was applied to the performance variables because it allows a more accurate comparison between different levels of difficulty by taking the target frequency into consideration.

Although subjects were engaged in the AIRDEF simulation at three levels of difficulty, behavioral performance measures were only collected during the medium (18 targets, medium tempo) and high (54 targets, high tempo) load conditions. This is because during the low load (baseline) condition, no targets were incoming.

The NSR measures were subjected to paired comparison t-tests (high load minus medium load difference) to determine if behavioral performance decreased as workload increased. Table 2 summarizes the results for both the visual and auditory attend conditions. In both the visual and auditory conditions subjects performed significantly better in the 18 targets condition than in the 54 targets condition (visual NSR difference, $t\ (63) = -3.89$, $p < .0002$; auditory NSR difference, $t\ (63) = -6.89$, $p < .0001$). There were no statistically significant behavioral differences between the visual and auditory attend conditions for any of the AIRDEF performance measures.

Physiological Analyses

Theory and previous results had indicated that the effects of interest, namely, ERP components related to workload effects, could be found using specific electrode sites and time windows. On this basis, two analyses were conducted. The first analysis was a repeated measures analysis of variance (ANOVA). The independent factors were workload and two electrode sites (P3 and F4). These sites were chosen because the frontal and parietal areas have been shown to be the most sensitive in attention and workload studies. P3 and F4 were specifically chosen to maximize separation across the scalp between these two brain regions. The dependent measure was the RMS amplitude computed over the window from 300-450 ms post-stimulus. This window usually corresponds to the third large positive component in the ERP and has been labeled the P300 component (Gopher & Donchin, 1986). It is this component that has been shown to be sensitive to workload-related changes in cognitive processing. We will refer to the 300-450 ms window RMS as the P300-RMS in what follows.

Table 2

A. NSR Performance Statistics for Attend Visual Conditions

Variable	Mean	SD
Medium load	62.95	12.86
High load	55.86	10.58

B. NSR Performance Statistics for Attend Auditory Conditions

Variable	Mean	SD
Medium load	65.29	10.42
High load	54.62	9.53

Separate ANOVAs were done for the visual and auditory ERPs. These are summarized in Table 3. Part A of Table 3 represents the ANOVA for the attend visual conditions. The main effect for workload was not significant $F(2, 128) = 0.16$, $p < .851$. There was a significant main effect for electrode site $F(1, 64) = 24.77$, $p < .000$, with site P3 having a larger P300-RMS amplitude across workload levels than site F4 (mean P300-RMS amplitude of 3.46 μV, and mean F4 amplitude of 2.96 μV).

Part B of Table 3 summarizes the ANOVA for the attend auditory conditions. There was a significant main effect for workload in this analysis $F(2, 128) = 3.50$, $p < .035$. There was a similar main effect for site in this condition as in the visual condition $F(1, 64) = 28.56$, $p < .000$, with site P3 having a larger P300-RMS amplitude across workload levels than site F4 (mean P300-RMS amplitude of 3.44 μV, and mean F4 amplitude of 2.97 μV).

Table 3

A. ANOVA Summary Table for Visual ERPs

Source	*SS*	*df*	*MS*	*F*	*p*
Workload (W)	0.76	2	0.38	0.16	.851
W × Subjects (Sb)	310.19	128	2.42		
Site (S)	24.93	1	24.93	24.77	.000
S × Sb	64.42	64	1.01		
W × S	0.20	2	0.10	0.17	.843
W × S × Sb	75.31	128	0.59		

B. ANOVA Summary Table for Auditory ERPs

Source	*SS*	*df*	*MS*	*F*	*p*
Workload (W)	14.14	2	7.07	3.50	.035
W × Subjects (Sb)	258.30	128	2.02		
Site (S)	21.11	1	21.11	28.56	.000
S × Sb	47.29	64	0.74		
W × S	0.39	2	0.20	0.29	.738
W × S × Sb	87.99	128	0.69		

The p values are the Greenhouse-Geisser corrected probabilities (BMDP Statistical Software, 1985).

The workload means and standard deviations for the attend auditory conditions ANOVA are shown in Table 4. Planned comparisons (planned statistical tests) were carried out on these means to test for statistical differences. A modified Bonferroni correction (Keppel, 1982) indicated that for these comparisons, the nominal F probability had to be $p < .03$ to be reported significant at the $p < .05$ level.

Table 5 summarizes the planned comparisons. As indicated, the only significant difference found was between the low load (baseline) and high load (54 targets) conditions, $F(1, 64) = 8.23, p < .006$.

The second set of analyses were similar to the above ANOVAs except an earlier time window in the ERPs was used as a covariate. Again, workload (three levels) and electrode sites P3 and F4 were the independent factors and P300-RMS (300-450 ms) was the dependent variable. The covariate was computed on slightly different time windows for the visual and auditory ERPs. For the visual ERPs, the window ranged from 125-227 ms post-stimulus and for the auditory ERPs it ranged from 94-195 ms post-stimulus. Each window was approximately 100 ms wide. An RMS amplitude measure was computed for those windows. These early windows have been referred to as the N1 component (Näätänen, & Picton, 1987). Different windows were chosen for the visual and auditory ERPs because the peak latency for the N1 component varies between the visual and auditory modalities (Näätänen, & Picton, 1987). We will refer to both of these windows as the N1-RMS in the following sections.

Separate two-way repeated measures analysis of covariance (ANCOVA) were carried out on the visual and auditory ERPs. The independent factors were workload (three levels) and electrode sites P3 and F4. The P300-RMS (300-450 ms) was the dependent variable, and the covariate was the N1-RMS for the visual and auditory ERPs (125-227 ms and 94-195 ms, respectively). The results are summarized in Table 6. Part A of Table 6 represents the ANCOVA for the visual ERPs. Again, the main effect for workload was not significant $F(2, 127) = 0.19, p < .822$. The N1-RMS was significantly related to the P300-RMS in the workload effect $F(1, 127) = 9.75, p < .002$. There was also a significant main effect for electrode site $F(1, 63) = 10.57, p < .002$. The adjusted mean for site P3 was larger than the adjusted mean for site F4 (see Table 7, part A). The N1-RMS was not significantly related to the P300-RMS for the site main effect $F(1, 63) = 0.78, p < .380$.

Part B of Table 6 summarizes the ANCOVA for the auditory ERPs. As with the ANOVA, the main effect for workload was significant $F(2, 127) = 5.92, p < .004$. The N1-RMS was significantly related to the P300-RMS for

Table 4

Workload Main Effect Means for Auditory ERPs

Difficulty Level	Mean	SD
Low load	3.45	1.43
Medium load	3.19	1.26
High load	2.98	1.13

Table 5

ANOVA Planned Comparisons Between Workload Levels for Auditory ERPs

Source	*SS*	*df*	*MS*	*F*	*p*
Low load vs medium load	4.21	1	4.21	1.73	.193
Error	155.86	64	2.44		
Low load vs high load	14.10	1	14.10	8.23	.006
Error	109.56	64	1.71		
Medium load vs high load	2.90	1	2.90	1.52	.222
Error	122.03	64	1.91		

Table 6

A. ANCOVA Summary Table for Visual ERPs

Source	*SS*	*df*	*MS*	*F*	*p*
Workload (W)	0.88	2	0.44	0.19	.822
Covariate (N1 Component)	22.11	1	22.11	9.75	.002
W × Subjects (Sb)	288.08	127	2.27		
Site (S)	10.68	1	10.68	10.57	.002
Covariate	0.79	1	0.79	0.78	.380
S × Sb	63.63	63	1.01		
W × S	0.27	2	0.13	0.23	.794
Covariate	1.54	1	1.54	2.65	.106
W × S × Sb	73.77	127	0.58		

B. ANCOVA Summary Table for Auditory ERPs

Source	*SS*	*df*	*MS*	*F*	*p*
Workload (W)	20.48	2	10.24	5.92	.004
Covariate (N1 Component)	38.56	1	38.56	22.29	.001
W × Subjects (Sb)	219.74	127	1.73		
Site (S)	24.72	1	24.72	39.89	.001
Covariate	8.25	1	8.25	13.30	.001
S × Sb	39.05	63	0.62		
W × S	0.47	2	0.23	0.37	.687
Covariate	7.53	1	7.53	11.89	.001
W × S × Sb	80.46	127	0.63		

this effect $F(1, 127) = 22.29, p < .001$. There was also a significant main effect for site $F(1, 63) = 39.89, p < .000$, with the N1-RMS being significantly related to it $F(1, 63) = 13.30, p < .001$. The adjusted means for P3 and F4 are shown in Table 7, part B. As with the unadjusted means, site P3 has a larger P300-RMS amplitude than site F4.

Planned comparisons were done on the adjusted workload main effect means for the auditory ERPs (see Table 8). As with the previous planned comparisons, family-wise error was controlled by the modified Bonferroni procedure (Keppel, 1982). Since the number of comparisons and degrees of freedom were identical, the nominal F probability needed to be $p < .03$ to be significant at the $p < .05$ level. There was a significant difference between the low load and high load conditions $F(1, 63) = 12.96, p < .001$, as was found when carrying out the ANCOVA planned comparisons. There was also a significant difference between the low load and medium load conditions $F(1, 63) = 6.72, p < .012$. The adjusted means decreased monotonically from the baseline to the high load condition (low load 3.52 µV, medium load 3.11 µV, and high load 2.98 µV). The N1-RMS was significantly related to all comparisons.

Analyses were conducted on the auditory ERPs in the ANCOVA to test for the assumptions of homogeneity of within-group regression lines and independence of the covariate from the independent treatment of workload. Results indicated that both of these assumptions were met, confirming the major assumptions for the ANCOVA (Huitema, 1980).

Table 7

A. Adjusted Means for Visual ERPs

Electrode Site	Mean	SD
P3	3.38	1.43
F4	3.05	1.21

B. Adjusted Means for Auditory ERPs

Electrode Site	Mean	SD
P3	3.46	1.33
F4	2.95	1.17

Table 8

ANCOVA Planned Comparisons Between Workload Levels for Auditory ERPs

Source	*SS*	*df*	*MS*	*F*	*p*
Low load vs medium load	12.81	1	12.81	6.72	.012
Covariate	35.80	1	35.80	18.79	.000
Error	120.06	63	1.91		
Low load vs high load	18.77	1	18.77	12.96	.001
Covariate	18.28	1	18.28	12.62	.001
Error	91.28	63	1.45		
Medium load vs high load	1.67	1	1.67	0.92	.340
Covariate	8.22	1	8.22	4.55	.037
Error	113.82	63	1.81		

DISCUSSION

The experimental results demonstrated that a specific ERP amplitude measure, P300-RMS at two electrode recording sites (P3 and F4), was sensitive to workload changes in the AIRDEF radar simulation. This finding agrees with studies reported in the literature that have used the P300 component of the ERP to measure workload (Gopher & Donchin, 1986). The results were enhanced by taking into consideration an early ERP amplitude measure, N1-RMS, which corresponds to the N1 component that has been related to selective attentional factors (Näätänen & Picton, 1987). Analyses of variance indicated different findings for the visual and auditory attend conditions. In the visual attend conditions, P300-RMS did not vary as a function of AIRDEF difficulty, as evidenced by a non-significant workload main effect $F(2, 128) = 0.16$, $p < .851$. This result is not easily explained by an undifferentiated capacity resource allocation model. As indexed by the normalized skill rating (NSR), in the attend visual conditions behavioral performance decreased between the 18 and 54 target conditions (t (63) = -3.89, $p <$ .002). Since behavioral performance decreased as a function of target frequency and tempo, the demands for cognitive resources to perform the task should have increased with this factor. In this case, fewer resources would have been left to process the visual oddball stimuli, resulting in a decrease in P300 amplitude. Similar physiological results were reported by Isreal, Wickens, Chesney, and Donchin (1980) using an air-traffic controller task: amplitude of the P300 component did not vary as a function of target frequency when the subject attended to intensity changes in a visual stimulus.

One possible explanation for the failure of the attended visual oddball stimuli to index changes in primary task resource demands makes use of a multiple-resource model of cognitive capacity (Wickens, 1980). Suppose

that visual resources can be exhausted independent of other (e.g., auditory) resources. Then the perceptual demands of the baseline (low workload) condition may have consumed a large proportion of disposable visual resources. This proportion may have been almost as large as that required in the higher workload conditions. Therefore, the net change in demands for visual resources may have been a small fraction of the total resources employed, which would lead to small changes in P300-RMS.

In the attend auditory conditions, P300-RMS decreased significantly as a function of task difficulty ($F(2, 128) = 3.50, p < .035$). Planned comparisons of these means demonstrated that the only reliable decrease was between the low load (baseline) condition and the high load (54 targets) condition (Table 5). Behavioral performance decreased between the medium load and high load target conditions when attending to the auditory stimuli ($t\,(63) = -6.89, p < .0001$). Similar results were reported by Isreal, Wickens, Chesney, and Donchin (1980) with respect to P300 amplitudes elicited by an auditory oddball stimulus while performing the air-traffic controller task. In the Isreal et al. study, reliable P300 amplitude differences were found between all difficulty levels. This included a count-only "baseline" condition and two levels of target frequency (four and eight). Reaction time data also decreased as a function of target frequency, behaviorally validating the workload effect.

The decreases observed in P300 amplitude are hypothesized to represent a decrease in the perceptual processing resources available to perform the oddball task with increases in AIRDEF task difficulty (Wickens, 1980). However, some artifactual and/or systematic possibilities other than perceptual processing resources need to be taken into consideration in validating the observed changes. Two other possibilities could account for the observed decreases in P300-RMS amplitude as a function of workload. First, since the amplitude of P300 is a function of signal averaging, increases in the trial-to-trial variability of the single epoch P300 amplitudes as a function of workload would reduce the averaged amplitude measure. Since single epoch P300 amplitudes were not measured in this study, this possibility can not be discounted. A second possibility is that the peak latency of the P300 component changed as a function of task workload, shifting the component in time. This would result in reduced amplitudes, since a constant time window was used across difficulty levels. Isreal, Wickens, Chesney, and Donchin (1980) found that there were significant peak latency changes in the P3 component with increasing difficulty levels in the air-traffic controller task when subjects attended to the auditory stimuli. Latency changes in the Isreal et al. study mirrored the amplitude measure changes and increased monotonically as a function of workload. Because a latency analysis was not carried out on the present data, the possibility exists that the observed reduction in P300-RMS amplitude between the low load and high load conditions of AIRDEF could be due at least in part to latency variability in the P300 component.

Based on possible interdependence of N1 and P300 (see Introduction), a repeated measures analysis of covariance (ANCOVA) was computed for the visual and auditory P300-RMS data. The N1-RMS was used as the covariate to factor out the variance attributed to attention so that a more reliable estimate of processing resources as measured by P300-RMS could be analyzed. For the visual data the results were not different from the ANOVA; however, the workload main effect for the auditory P300-RMS data was enhanced. Planned comparisons on the adjusted workload mean values demonstrated reliable differences between the low load condition and the medium load condition as well as between the low load and high load conditions (Table 8).

The N1-RMS was insensitive to workload-related processing as it did not change significantly with that effect ($F(2, 128) = 2.56, p < .082$.) These results contradict the notion that N1 and P300 are independent (Näätänen & Picton, 1987). Most of the work in the area of selective attention and N1 has not employed the same methodologies as used with P300 and mental capacity assessment. These differences could account for the prior lack of evidence supporting a relationship between N1 and P300. Another reason for this has been the lack of interaction between researchers involved in N1 and P300 research.

Our results confirmed previous findings with visual and auditory oddball stimuli as secondary tasks. The addition of early processing resources to the model appears to increase its predictive power, but this is a post hoc explanation that needs to be confirmed by future studies. Even with this addition, the auditory findings are not fully explained. Clearly the difficulty of the primary task increased between the three levels of workload manipulation. If the auditory attentional pool was undisturbed by the primary task, then maximum channel information would have been passed to later evaluation processes. It may be that attentional resources are not completely separate between stimulus modalities, especially in the central processing stages (Wickens, 1980). There may also be a limit at higher ends of difficulty manipulation in relation to physiological responses. In fact, some researchers have suggested that the physiological information contributes maximum prediction in studying workload-related processes at the lower levels of difficulty (O'Donnell & Eggemeier, 1986). This makes sense in light of the experimental evidence.

FUTURE APPLICATIONS

Future studies can take these variate sets and maximize their unique contributions through multiple regression and canonical correlation analysis, for example, in the prediction of on-job performance from traditional training grades. If simulation performance scores are used to predict radar operator job performance criteria, this prediction would be augmented by optimally combining the physiological and behavioral variates to predict these simple or composite criterion measures.

Our results support the resource allocation model of individual cognitive processes. It is recommended that future studies explore several workload levels covering the lower difficulty spectrum and other means of varying workload. This should be done in order to try to maximize the contribution of the physiological variables sensitive to these changes.

The results also suggest the possibility of on-line monitoring of cognitive resources using ERPs. Such monitoring could aid in maintaining an estimate of operator effectiveness in critical tasks.

REFERENCES

BMDP Statistical Software. (1985). *P2V--Analysis of variance and covariance including repeated measures.* Berkeley. University of California Press.

Brown, I. D. (1978). Dual task methods of assessing work-load. *Ergonomics, 21*, 221-224.

Callaway, E. (1975). *Brain electrical potentials and individual psychological differences.* New York: Grune & Stratton.

Callaway, E., Halliday, R., & Herning, R. I. (1983). A comparison of methods for measuring event-related potentials. *Electroencephalography and Clinical Neurophysiology, 55*, 227-232.

Defayolle, M., Dinard, J. P., & Gentil, M. T. (1971). Averaged evoked potentials in relation to attitude, mental load and intelligence. In W. T. Singleton, J. G. Fox, & D. Whitfield (Eds.), *Measurement of man at work.* New York: Van Nostrand Reinhold.

Donchin, E., Kramer, A. F., & Wickens, C. D. (1982). Probing the cognitive infra-structure with event-related brain potentials. In M. C. Frazier & R. B. Crombie (Eds.), *Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics* (AFFTC-JR-82-5), pp. 371-387.

Donchin, E., Kramer, A. F., & Wickens, C. D. (1986). Applications of brain event-related potentials to problems in Engineering Psychology. In M. G. H. Coles, E. Donchin, & S. Porges (Eds.). *Psychophysiology: Systems, processes, and applications.* Middleton, NJ: Till & Till.

Duncan-Johnson, C. C., & Donchin, E. (1979). The time constant in P300 recording. *Psychophysiology, 16(1)*, 53-55.

Fishman, J. A. (1958). Unsolved criterion problems in the selection of college students. *Harvard Educational Review, 28(4)*, 340-349.

Ghiselli, E. E. (1966). *The validity of occupational aptitude tests.* New York: John Wiley.

Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26*, 461-477.

Gopher, D., & Donchin, E. (1986). Workload--An examination of the concept. In K. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. II.* New York: John Wiley.

Heffley, E., Wickens, C. D., & Donchin, E. (1978). Intramodality selective attention and P300--Reexamination in a visual monitoring task. *Psychophysiology, 15*, 269-290.

Huitema, B. E. (1980). *The analysis of covariance and alternatives.* New York: John Wiley.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology, 17(3)*, 259-273.

Isreal, J. B., Wickens, C. D., Chesney, G. L., & Donchin, E. (1980). The event-related brain potential as an index of display monitoring workload. *Human Factors, 22(2)*, 211-224.

Jasper, H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10*, 211-224.

Johnson, R., Jr. (1978). P300: A model of the variables controlling its amplitude. *Annals: New York Academy of Sciences, 112*, 223-229.

Johnson, R., Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology, 23(4)*, 367-384.

Johnson, R., Jr., & Donchin, E. (1985). Second thoughts: Multiple P300s elicited by a single stimulus. *Psychophysiology, 22(2)*, 182-194.

Karis, D., Coles, M. G. H., & Donchin, E. (1984). *The use of event related brain potentials for prediction of performance quality in pilots of advanced tactical jet aircraft* (Tech. Rep. CPL 84-2/F33615-82-K-0611). Champaign: CPL, University of Illinois.

Kelly, R. T., Greitzer, F. L., & Hershman, R. L. (July 1981). *Air Defense: A computer game for research in human performance* (NPRDC Tech. Rep. 81-15). San Diego: Navy Personnel Research and Development Center.

Keppel, G. (1982). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice Hall.

Kramer, A. F., Wickens, C. D., & Donchin, E. (1986). Processing of stimulus properties: Evidence for dual-task integrality. *Journal of Experimental Psychology: Human Perception and Performance, 11(4)*, 393-408.

Kramer, A. F., Wickens, C. D., Vanasse, L., Heffley, E. F., & Donchin, E. (1981). Primary and secondary task analysis of step tracking: An event related potentials approach. In R. C. Sugarman (Ed.), *Proceedings of the Twenty-Fifth Annual Meeting of the Human Factors Society*, Rochester, NY.

Lewis, G. W. (1979). Visual event related potentials of pilots and navigators. In D. Lehmann & E. Callaway (Eds.), *Human evoked potentials. Applications and problems*. New York: Plenum Press.

Lewis, G. W. (1983a). Event related brain electrical and magnetic activity: Toward predicting on-job performance. *International Journal of Neuroscience, 18*, 159-182.

Lewis, G. W. (1983b). *Bioelectric predictors of personnel performance: A review of relevant research at the Navy Personnel Research and Development Center* (NPRDC Tech. Rep. 84-3). San Diego: Navy Personnel Research and Development Center.

Lewis, G. W. (1984). Temporal stability of multichannel, multimodal ERP recordings. *International Journal of Neuroscience, 25*, 131-144.

Lewis, G. W., & Froning, J. N. (1981). Sensory interaction, brain activity, and reading ability in young adults. *International Journal of Neuroscience, 15*, 129-140.

Lewis, G. W., & Rimland, B. (1979). *Hemispheric asymmetry as related to pilot and radar intercept officer performance* (NPRDC Tech. Rep. 79-13). San Diego: Navy Personnel Research and Development Center.

Lewis, G. W., & Rimland, B. (1980). *Psychobiological measures as predictors of sonar operator performance* (NPRDC Tech. Rep. 80-26). San Diego: Navy Personnel Research and Development Center.

Lewis, G. W., Rimland, B., & Callaway, E. (1977). *Psychobiological predictors of success in a Navy remedial reading program* (NPRDC Tech. Rep. 77-13). San Diego: Navy Personnel Research and Development Center.

Lewis, G. W., & Sorenson, R. C. (October 1987). *Evoked brain activity and personnel performance* (NPRDC Tech. Rep. 88-1). San Diego: Navy Personnel Research and Development Center.

Lewis, G. W., Trejo, L. J., Blackburn, M. R., & Blankenship, M. H. (1986). Neuroelectric and neuromagnetic recordings: Possible new predictors of on-job performance. In G. E. Lee (Ed.), *Proceedings: Psychology in the Department of Defense. Tenth Symposium* (USAFA TR 86-1). Colorado Springs: U. S. Air Force Academy.

Lindholm, E., Cheatham, C., & Koriath, J. (April 1984). *Physiological assessment of aircraft pilot workload in simulated landing and simulated hostile threat environments* (AFHRL TR 83-49). Brooks Air Force Base, Texas.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24(4)*, 375-424.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7*, 44-64.

O'Donnell, R. D., & Eggemeier, F. (1986). Workload assessment methodology. In K. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. II.* New York: John Wiley.

Papanicolaou, A. C., & Johnstone, J. (1984). Probe evoked potentials: Theory, method and applications. *International Journal of Neuroscience, 24*, 107-131.

Pritchard, W. (1981). Psychophysiology of P300. *Psychophysiology, 89(3)*, 506-540.

Rolfe, J. M. (1971). The secondary task as a measure of mental workload. In W. T. Singleton, J. G. Fox, & D. Whitfield (Eds.), *Measurement of man at work.* London: Taylor and Francis.

Sirevaag, E., Kramer, A. F., Coles, M. G. H., & Donchin, E. (in press). Resource reciprocity: An event-related brain potentials analysis. *Human Factors.*

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked potentials correlates of stimulus uncertainty. *Science, 150*, 1187-1188.

Trejo, L. J. (April 1986). *Brain activity during tactical decision-making: I. Hypotheses and experimental design* (HFOSL Tech. Note 71-86-6). San Diego: Navy Personnel Research and Development Center.

Trejo, L. J., Lewis, G. W., & Blankenship, M. H. (1987). *Brain activity during tactical decision-making: II. Probe-evoked potentials and workload* (NPRDC Tech. Note 88-12). San Diego: Navy Personnel Research and Development Center.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance. VIII.* New Jersey: Lawrence Erlbaum Associates.

Wickens, C. D., Isreal, J., & Donchin, E. (1977). The event related cortical potential as an index of task workload. *Proceedings of the Twenty-first Annual Meeting of the Human Factors Society*, Santa Monica, CA, pp. 282-286.

DISTRIBUTION LIST

Office of the Chief of Naval Operations (OP-933D4)
Deputy Under Secretary of Defense for Research and Engineering
Assistant for Manpower Personnel and Training Research and Development (OP-01B2)
Director, Command Surface Warfare Manpower and Training Requirements (OP-39)
Naval Training Systems Center, Technical Library (5)
Head, Cognitive and Neurosciences (OCNR-1142)
Life Sciences Directorate (OCNR-114)
Perceptual Science (OCNR-1142PS)
Cognitive and Decision Science (OCNR-1142CS)
Technology Area Manager, Office of Naval Technology (Code 222), (Code 223)
Commanding Officer, Naval Aerospace Medical Research Laboratory, Pensacola, FL (Code 031)
Commanding Officer, Naval Health Sciences Education and Training Command, Bethesda, MD
Naval Medical Research and Development Command (Code 40)
Naval Medical Command (02D)
Naval Biodynamics Laboratory
Commanding Officer, Naval Hospital, Portsmouth, VA (Medical Library)
Army Research Institute (PERI-SB)
Headquarters, Marine Corps (Code MA)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base, TX
Superintendent, Naval Postgraduate School
Director of Research, U.S. Naval Academy
Program Manager, Manpower Research and Advisory Service, Smithsonian Institute
Center for Naval Analyses, Acquisitions Unit
Defense Technical Information Center (DTIC) (2)